DOCUMENT RESUME

ED 053 194 TM 000 724

AUTHOR Freiberg, H. Jerome
TITLE The Effects of Ability Grouping on Interaction in the Classroom.
PUB DATE 70
NOTE 15p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Ability Grouping, *Classroom Communication, Computer Oriented Programs, Discriminant Analysis, History Instruction, Homogeneous Grouping, Interaction, Interaction Process Analysis, *Pilot Projects, Reinforcement, Secondary Schools, Student Motivation, *Student Teacher Relationship, *Teacher Behavior
IDENTIFIERS Flanders Interaction Analysis

ABSTRACT

This study represents classroom observations of a secondary school which employs ability grouping for grades seven through twelve. The effects of ability grouping on the interaction in the classroom was measured with use of Flanders Interaction Analysis. Two history teachers with identical grades and phases were used in this study which began in the Fall of 1970 and ended that January. Each teacher's ability groups were observed twice during the semester, once at the beginning of the year and once near the end of the term. Darwin's Chi Square and graphic representations of frequency counts were used to measure differences between the two teachers and within their own ability groups. Preliminary results show that the upper phases tend to receive more empathy, praise and use of their ideas and less direction and criticism than the lower phases according to the Flanders Interaction Analysis system. The two teachers also tend to teach their lower phases in a similar manner. (Author/AG)

ED053194

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Paper Presented by H. Jerome Freiberg

on "The Effects of Ability Grouping on the Interaction in the Classroom"

This paper represents an abstract of a pilot study conducted at a High School in Western Massachusetts during the fall of 1970.

The pilot study is a product of my observations as a supervisor for the University of Massachusetts during the past two years. I observed two female teachers (A and B), both History teachers in the seventh and eighth grades. The teachers were selected for their diversity of ability groups. Teacher A taught three ability (or phase) levels (IV, III, II, and I) while Teacher B taught three ability levels (IV, III, and II). The students were placed into ability groups according to their I.Q. and the previous semester's grades. Phase IV represents the highest ability group and Phase I the lowest.

I used Flanders Interaction Analysis as my observational tool for this study. Flanders Interaction Analysis enables me to quantify teaching behaviors so I may produce "relatively unbiased judgements" on their teaching patterns in the classroom.

TM 000 724

With the use of Flanders Interaction Analysis and its indirectness/directness ratios I am attempting to observe behavioral interactions in twelve teaching patterns:

1) The percentage of teacher talk as compared with student talk.
2) The general pattern of directness or indirectness in the classroom.
3) Various types of motivation used by the teacher.
4) The differences in motivating behavior and controlling behavior.
5) The degree of reinforcement the teacher employes after the student responds.
6) The type of responses the teacher reinforces.

7) The extent to which the class is motivated without the use of content.
8) The time lapse between the first and second teacher question.
9) Student reliance upon the teacher as the focal point in classroom discussion.
10) The types of questions the teacher asks the students.
11) The extent to which the teacher uses the students' ideas.
12) The general teaching pattern of the classroom.

If the data does not support the assumptions I have made in this paper I may conclude that: 1) I am using an inappropriate research tool; 2) I need more data; 3) I should observe another group of teachers; 4) I need to approach the problem from another direction; 5) My observations were misleading and my assumptions were incorrect.

The observed changes will be viewed between the different phases of each teacher, between the first and second observations and between the two teachers. By using the teacher as the focal point of this study, I hope to analyze the behavioral changes in teaching patterns as they interact with various homogeneous groups of students. Finally the study should provide a basis for the acceptance or rejection of my hypothesis of the negative effects of ability grouping on interaction in the classroom.

I will calculate differences between and within teachers phases graphically, using percentages of frequency counts and Darwin's Chi Square. The problems of analysis are discussed in greater detail in the preceding pages.

I observed each teacher's class once in the beginning of the semester and then repeated the observation near the end of the semester. The observations were completed over a period of a week during the first and second observations. To keep the variables of content somewhat equal, I only viewed the class the teachers designated as review sessions. The teachers were not chosen at random for a number of reasons: Firstly, there were only a few teachers willing to participate in the study and they were in other content areas and grade levels. Secondly, I needed the broadest possible range of phases to analyze the reaction of the teachers' behavior as different groups of students entered the teachers' classes.

I recorded the observations myself via a tape recorder which I used in the rear of the room. I spent a few days in the classes that were not labeled as review to give the students the opportunity to get accustomed to me and my tape recorder. I rated the tapes myself, which brings out the problem of rater-reliability rather than inter-rater-reliability.

After completing the observations, I fed my tallies onto computer cards and read the data into a program developed by Lawrence Wightman (1970) for a ten category Flanders Interaction Analysis (see appendix A for an explanation of Flanders Interaction Analysis and its uses) scale with a ten by ten matrix. The output will be discussed in the analysis section of this paper.

The use of Flanders Interaction Analysis presents some problems in that, "the exact usefulness of IA in predicting pupil achievement remains to be seen." (Rosenshine, 1970.) The use of Flanders Interaction Analysis (Flanders, 1965) assumes that an indirect teacher increases the achievement level of her students, and that direct teacher's behavior limits the pupil's freedom, as defined by the ten categories in his system. In Rosenshine's review of the Flanders Interaction Analysis research and in my own reading there appears to be a deficiency of valid statistical analyses. The validity errors range from using the number of students for the unit of analysis rather than the teachers which were the sampling unit. (Rosenshine, et. al.) In many of the other studies less than five teachers were used in the study, with the authors generalizing past the bounds for a study intended for a limited population. In my review of Journal of Research and Development of Systematic Observation (Fall, 1970), the systems which were explored used simple t-tests for expanded Flanders categories without considering adjustment of t-tests. The simple mean comparisons ignore the fact that a closed matrix system creates the problems of didactic variables; that is, the cell which receives a specific number of tallies effects the outcome of the remaining categories. There is also the problem that the relationships between Flanders categories and student achievement may not represent a linear relationship (Soar, 1968). The use of Chi-Square statistics within and between teachers compounds the problems of didactic variables and the comparison of dependent variables. Some systems get around

the problems by using frequency counts of specific variables, leaving the system open, but this process does not necessarily represent a more appropriate approach.

Since Flanders developed his interaction system (Flanders, 1960), some 400 observational systems have been created, but fewer than 15 have been developed and used to relate classroom behaviors with student growth (Rosenshine and Furst, 1971) and many of these systems face the problems of inadequate or inappropriate statistical analyses (Greenberg, 1970). I am approaching the research problem of ability grouping with a tool that will need special and perhaps newly developed statistical analysis. Since most of the other observational systems in use suffer from similar maladies, I feel the use of the Flanders System is still justifiable.

In my position paper I discussed the twelve behaviors of variables I intended to measure, but realizing the problems of statistical analysis I decided to approach the analyses from a more total approach which would include some but not all the variables. Using the total matrix for each of the classes of Teacher A and B I am presented with eight matricies for Teacher A and six matricies for Teacher B (see appendix C). Keeping in mind the previous discussion on matrix analysis and its problems in relations to the Flanders System, I decided to use the Darwin's Chi-Square in my somewhat contrived "goodness of fit" between teachers and within the teachers. My first analysis tests the differences or similarities between the first observation (O_1) and the second observation (O_2) along identical phases (see Table 1). Since the time (two months) elapsed between O_1 and O_2 I was interested to see if the teachers' (matricies) interaction with their classes were different or similar. For Teacher A only Phase IV displayed a difference at the .005 level using Darwins Chi-Square, the other three phases of Teacher A did not develop any significant differences. Teacher B showed significant difference with Phase II with the other two phases not showing a significant at the .10 level. It would seem that according to my previous limitations, I could make the reserved generalization that over a period of time these two teachers tend to interact with identical phases (see Table 1), in a similar manner. Between

Teacher A and B Phases IV and III showed significant differences in classroom interaction at the .005 level, with Phase II of both teachers developing no significant differences. Using the reservations of the above analysis I would perceive that Teacher A and B interact differently with their classes except on the lower ability level of Phase II. Since Teacher B did not have a Phase I class I was unable to measure similarities or differences on that particular phase level.

One of the important variables to be tested was the change in teacher-class interaction which occurred as the teacher interacted with different phase levels. Using Darwin's Chi-Square, I compared Teacher A's first observations of Phase IV with Phase III, Phase III with Phase II and Phase II with Phase I. Significant differences were obtained in all four phase levels (see Table 2). I repeated the test using the same statisitcal techniques with Teacher A's second set of observations. Once again I was comparing the total matricies for goodness of fit with the matricies comprising the other phases. Since I am considering a level of confidence of .05 as a significant bounds for my study, the differences between Phase IV and III approach very closely to a significant difference. The measurement between Phase III and II showed no significant difference while comparison between Phase II and I achieved significance beyond the .005 level (see Table 2). Other comparisons which I have not completed will be a measure of difference between Phase IV and I and a pairing of Phases of IV and III and measuring them against Phases II and I. Of the six possible measures between individual matricies, four showed significant difference, one approached significance and comparison developed no significant difference. I would conclude from these results that Teacher A tended to interact with her phases differently while interacting with identical phases in a similar manner.

My next statistical comparison follows the path of the above analysis by doing a cross comparison of O_1 and O_2 of Phases IV and III, Phases II and III, and Phases II and I (see Table 2). All comparisons either achieved or approached significance except for the comparison of Phase III (O_1) with Phase II (O_2). The greatest significance was achieved with the comparison of the two lower phases.

The Program Interact output develops a matrix for each class and calculates the number and percent of tallies in each cell and category, it also generates I/d ratios and elementary statistical analyses (see appendix A). Each of the ten Flanders categories represents defined interaction. A comparison within teachers will give some idea as to the trends different phases within the same teacher will show via Flanders ten categories (see Tables 3 and 4).

Table 4 represents the mean of the percents between O_1 and O_2 for each phase, and the graph presents a picture of how these percents relate to the ten Flanders categories. Appendix A will give you some idea as to the meanings for each of the ten categories, but basically the first three categories represent an indirect approach to teaching, with categories 6 and 7 representing, giving directions and criticism respectively. In the first three categories it appears the distributions seem fairly close except for Phase III which fluctuates from each of the other phases in its file. Teacher B differences are greatest between Phase IV and II, with the greatest amount of indirect interaction given to her upper phase (slightly high indirect, No. 1, No. 2, and No. 3 and a low percentage of tallies in No. 6 and No. 7) and a higher percent of tallies for phase 2 in categories No. 6 and No. 7. In the Flanders system there is an interaction 9-10-9 within categories the amount of student-student interaction. Nearly 40 percent of Phase IV and 30 percent of Phase II tallies in categories No. 9 and No. 10 represent this student-student interaction. As a side note, Teacher B was not rehired because of her lack of "control" of her classes and especially her lower phase class. I cannot infer a cause and effect relationship, but it does seem to be an "interesting" relationship.

Table 3 represents the mean of the percents between O_1 and O_2 for each phase, and the upper phases were grouped into one by taking the mean percents of each phase. The same procedures were followed for the lower phases. It would appear that the differences are more distinct as the upper phases received a higher percentage of tallies in the indirect categories and a lower percentage of tallies in the direct area.

The analyses I have described in this paper have not met the variables for change I had described in the previous pages (p.1), but they will be measured at a later date. Presently I have the information available in the form of I/D ratios and percents to make some comparisons, but I feel a global approach to the data will be sufficient until I find statistical analyses which will not violate the assumptions for which I am testing.

CONCLUSION OF THE PILOT STUDY

My purpose in developing and testing a pilot study was to "further refine and define my hypothesis and afford me the opportunity to correct errors before I expend substantial quantities of time." My pilot project has de-fuzzed some concepts while creating many more fuzzy concepts. Although I planned to limit the amount of time by developing a pilot, I have realized that enormous amounts of time and energy that has gone into my pilot study, which gives me some inclinations as to the expenditures of time needed for a larger study with greater Teacher and observational N's.

The pilot has refined my thinking as to what I am specifically attempting to answer and by what means I will pursue these answers. Such changes as focusing in on teacher interaction rather than teacher behavior will tighten my observations to those behaviors which deal with verbal classroom interaction and more specifically those which are defined by the Flanders category scale. The concept of the labeling effect is important but not a necessary part of my study. I have come to realize that ability grouping must be seen in the context of specific criteria and process. For example a system that bases its ability groups or phases on I. Q. tests and grade point averages without allowing the flexibility for a student to move in or out of a particular phase is certainly different from ability grouping which allows the student to move freely between phases without being "tracked" to one particular phase. If the criteria used to develop the ability group is different between schools, then additional variables are interacting with the general term "ability grouping."

I am using Flanders Interaction Analysis to test the interaction in the class, with the realization that special statistical analyses must be developed to deal with the problems of a closed observational system and matrix analysis. Through my research of the literature I have discovered a divergence of opinion regarding the relationship between achievement and direct/indirect teachers (Flanders, 1965, and Rosenshine and Furst, 1971). I have refined my hypothesis to incorporate the following:

Hypotheses

The hypotheses will be stated in general terms with each of the variables listed below being tested as part of the general hypotheses.

1) The teachers of identical phases will interact with their classes in a similar manner during the four observed class meetings. (See Table I, Phase II $O_1 - O_2$)

2) Two teachers will interact differently when the upper phases (e.g. Phases IV and III) are compared between teachers the theachers will interact in a similar manner when comparing their lower phases.

3) Teachers will interact differently with different ability groups (see Table 2, Teacher A, $O_1 - O_2$).

4) The greatest difference in the teacher's classes will occur between her highest and lowest ability group.

5) (A) The teacher's upper ability groups will receive a greater percentage of tallies on the Flanders Interaction Scale in categories Nos. 1, 2 and 3 and the least amount of tallies in categories 6 and 7.

   (B) The teacher's lower ability groups will receive a greater percentage of tallies on the Flanders Interaction Scale in categories No. 6 and 7, and the least percentage of tallies in categories 1, 2 and 3.

6) Teachers will be more direct, according to Flanders system, with lower ability groups and less direct with upper ability groups.

Variables to be measured between teachers and phases with regard to Flanders Interaction Analysis

1) Percentage of teacher talk as compared with student talk.

2) Various types of motivation used by the teacher.

3) The differences in motivating and controlling behaviors.

4) The type of responses the teacher reinforces.

5) The extent to which the class is motivated without the use of content.

6) Student reliance upon the teacher as the focal point in classroom discussion.

From the pilot study I will attempt to clarify and develop my position paper into a workable proposal for researching and studying the effects of teacher interaction with similar and different ability groups.

TABLE I

Comparision of Phases[1] Taught by Teacher A. and B. Using Darwin's Chi-Square Analysis[2]

Teacher A.			Teacher B.		
First Observation IA.	Significance	Second Observation 2A.	First Observation IB.	Significance[3]	Second Observation 2B.
PhaseIV	.005	PhaseIV	PhaseIV	Not sig. at .10 level	PhaseIV
PhaseIII	Not sig. at .10 level	PhaseIII	PhaseIII	Not sig. at .10 level	PhaseIII
PhaseII	Not sig. at .10 level	PhaseII	PhaseII	.005	PhaseII
PhaseI	Not sig. at .10 level	PhaseI	No Comparision	X	X

[1]Phase IV represents the top ability group, Phase III the average ability group, Phase II the lower ability group and Phase I the lowest ability group. Phases IV, III, II and I average a class size of thirty students, while Phase I usually limits its class size to ten students

[2]Darwin's Chi-Square is utilized by Dr. Lawrence Wightman as his statistical analyses for his Program Interact, which I used to run and analyze my data.

[3]The significance between Teacher A. and B. represents the comparing of matrix IA. and 2A. of Teacher A., with IB. and 2B. of Teacher B.

TABLE II

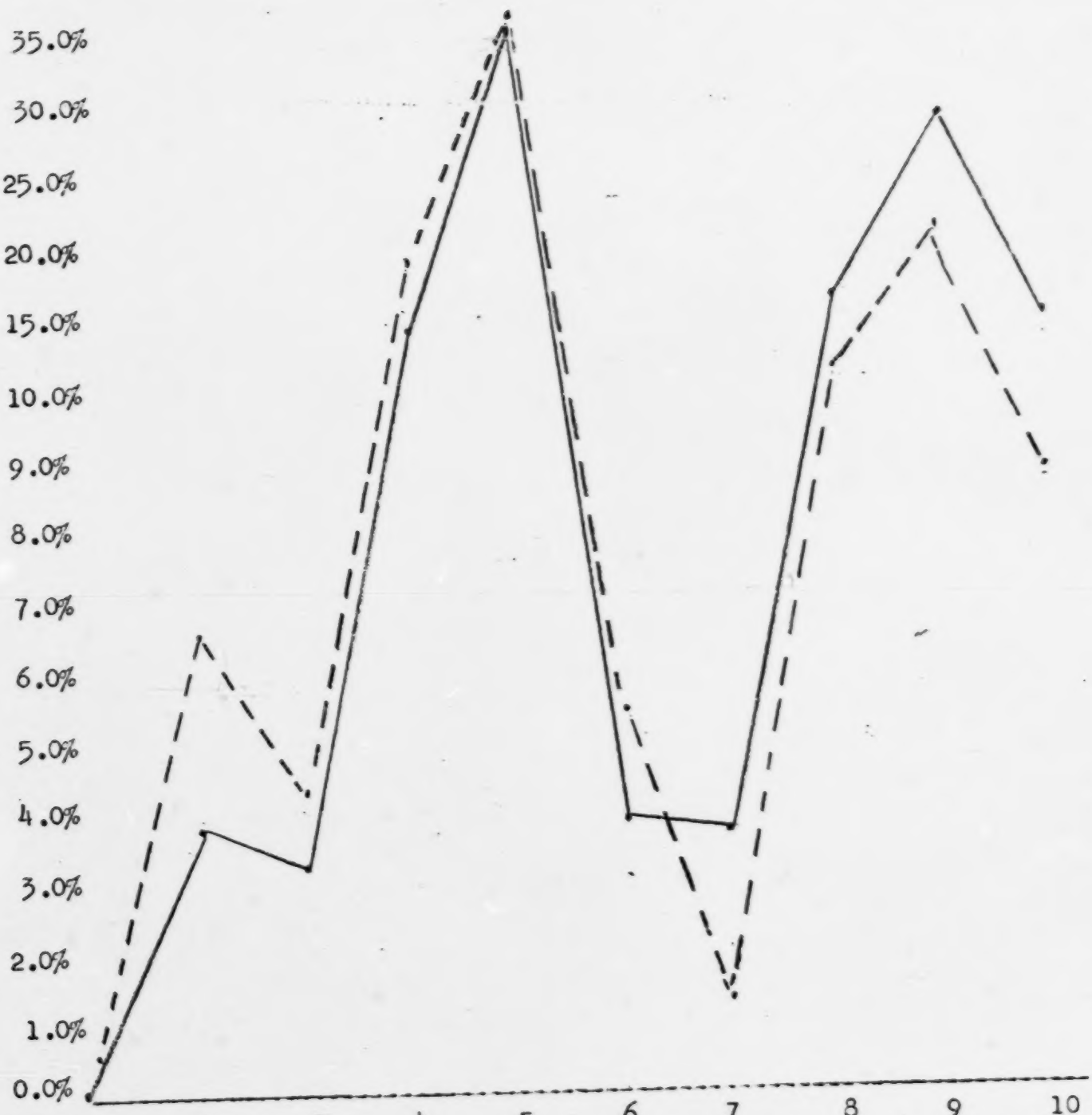
Teacher A.

First Observation	Second Observation
IA.	2A.
Phase IV	Phase IV
(sig.) .0.0-.025	(sig.) .05-.10
Phase III	Phase III
(sig.) .025-.05	(not sig.) at .10 level
Phase II	Phase II
(sig.) beyond .005	(sig.) beyond .005
Phase I	Phase I

Cross Comparison of Phases Taught by Teacher A Using Darwin's Chi-Square

Phase IV (IA.)	Significant between .025 - .05	Phase III (2A.)
Phase IV (2A.)	Significant between .05 - .10	Phase III (IA.)
Phase III (IA.)	Not significant at .10 level	Phase II (2A.)
Phase III (2A.)	Significant between .05 - .10	Phase II (IA.)
Phase II (IA.)	Significant beyond .005 level	Phase I (2A.)
Phase II (2A.)	Significant beyond .005 level	Phase I (IA.)

TABLE 3
Teacher A.

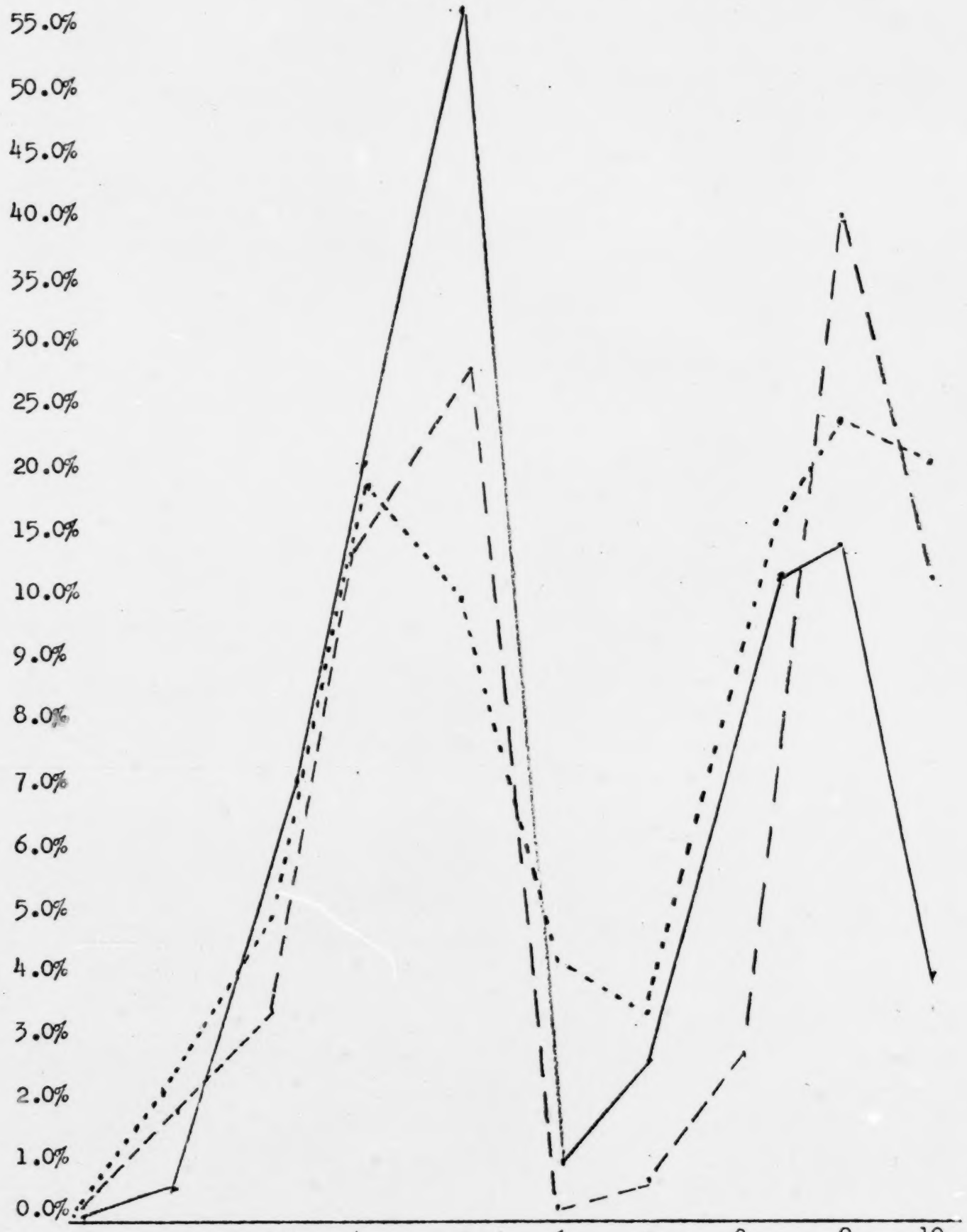


Flanders Catagories	1	2	3	4	5	6	7	8	9	10
$\bar{X}$ Phases IV and III	.61%	6.52%	4.31%	18.0%	35.98%	5.48%	1.38%	12.75%	20.51%	8.82%
$\bar{X}$ Phases II and I	0.0%	3.76%	3.35%	13.74%	35.83%	3.79%	3.53%	15.66%	30.44%	14.54%

(Upper Phases) Phases IV and III ----------

Teacher A. (Lower Phases) Phases II and I ——————

TABLE 4
Teacher B.



Flanders Catagories	1	2	3	4	5	6	7	8	9	10
Phase IV	.16%	1.88%	3.46%	13.62%	27.09%	.16%	.67%	2.62%	39.81%	10.31%
Phase III	0.0%	.54%	7.15%	20.39%	54.20%	.94%	2.50%	11.49%	13.55%	3.79%
Phase II	0.0%	2.08%	4.73%	18.29%	9.44	4.12%	3.29%	15.12%	22.96%	19.94%

Phase IV ----------
Phase III——————
Phase II ••••••••••••••

References

Beez,Victor W.,"Influence of biased psychological reports on teacher behavior and pupil performance." Miles,Mathew W., and Charters, W.W.Jr.,eds.,Learning in Social Settings(1970),pp.328-336.

Borg,Walter R.,"Ability Grouping in The Puplic Schools" Dembar Educational Research Services(1965).

Cohen,Ruth, and Isenberg,"Grouping For Instruction"(1967),ERIC,Ed.056829.

Ekstrom,Ruth B. Experimental Studies of Homogeneous Grouping,A Review of Literature(Princeton,N.J. : Educational Testing Services,1959).

Flanders,Ned A. "Teacher Influence,Pupil Attitudes, and Achievement." U.S. Office of Education Cooperative Research Project No.397. Minneapolis: University of Minnesota,1960. Unpublished Material

Flanders,Ned A. Teacher Influence,Pupil Attitudes, and Achievement. Cooperative Research Monograph No.12. Washington,D.C.: U.S. Department of Health, Educationa, and Welfare. Office of Education,1965.

Flowers,Charles, "Study of Ability Grouping"(1966) eds. Rosenthal,Robert, and Jacobson,Lenore,Pygmalion in the Classroom,Holt, Rinehart and Winston, New York,1968.

Greenberg, Selma, Selected Studies of Classroom Teaching: A comparative Analysis, International Textbook Company, Scranton Pennsylvania(1970).

Gumpert, and Gumpert,Urban Review,1968,3(1),21-25.

Jose,Norma Jean,"Teacher-Pupil Interaction as it Relates to Attempted Changes in Teacher Expectancy of Academic Ability and Achievement." Unpublished Dissertation(1970).

Norwin,William W.,"Some Practices in Grouping" Childhood Education,December 1968, Volume 45,Number 4 pp.189-196.

Rosenshine,Barak,"Interaction Analysis: A Tardy Comment," Phi Delta Kappan, Volume LI,Number 8,April,1970.

Rosenshine,Barak, and Furst,Norma,"Current and Future Research On Teacher Performance Criteria." ed. B.O.Smith,Research on Teacher Education: A symposium. Englewood Cliffs, New Jersey:Prentice-Hall,1971

Schrank,Wilburn R."The Labeling Effects of Ability Grouping," Journal of Educational Research,62:51,52,October 1968.

Schrank,Wilburn R. "AFurther Study of the Labeling Effect of Ability Grouping," Journal of Educational Research,63:8,pp. 358-361,1970.

Soar, R.S.,"Optimum Teacher-Pupil Interaction for Pupil Growth,"Educational Leadership, December,1968,pp.275-98.

References

"The Amherst Ungraded Secondary Schools,An Evaluation Report," U.S. Office Of Education, Department of Health,Education, and Welfare. E.S.E.A. Title III OEG-1-6-066873-1642,Volume I.

Thorndike,E.L., American Educational Research Journal,1968,5(4),708-711.

Wightman,Lawrence,"Achievement as a Function of Interactions Between Student Characteristics and Teacher Behaviors," Unpublished Thesis,Cornell University, September,1970.